CLAIMS

What is claimed is:

1   1. A method for efficiently batching transactions in a computerized system, comprising
2   the steps of:
3   receiving a transaction at the computerized system;
4   writing the received transaction data to a data buffer;
5   if the received transaction is an only transaction in the computerized
6   system, then flushing the data buffer to a stable storage device;
7   if the received transaction is not the only transaction in the computerized
8   system, then determining whether a maximum response time will be met for an oldest
9   transaction waiting for flush in the computerized system if the data buffer were flushed to
10 the stable storage device immediately;
11   if the maximum response time will not be met, then flushing the data buffer to the
12 stable storage device;
13   if the maximum response time will be met, then determining whether a rate of
14 data of incoming transactions is higher than a rate of flushing data from the data buffer to
15 the stable storage device if flushed immediately;
16   if the rate of data of incoming transactions is lower than the rate of flushing data
17 from the data buffer to the stable storage device if flushed immediately, then flushing the
18 data buffer to the stable storage device;
19   if the rate of incoming data is higher than the rate of flushing data from the data
20 buffer to the stable storage device if flushed immediately, then waiting to receive another
21 transaction; and
22   updating a flush performance history after each flush of the data buffer to the
23 stable storage device.

1   2. The method of claim 1 wherein the step of determining whether the maximum
2   response time will be met, comprises the steps of:

3   determining a time queued of the oldest transaction waiting for flush in the

4   computerized system;

5   determining an amount of data in the data buffer;

6   using the flush performance history to compute a time duration to flush the data

7   buffer to the stable storage device based on the amount of data in the data buffer;

8   calculating an expected delay in flushing the oldest transaction to the stable

9   storage device by subtracting the time when the oldest transaction announced that it was

10   waiting for flush from a current time and adding the time to flush the data buffer to the

11   stable storage device;

12   comparing the expected delay to a maximum response value;

13   if the expected delay is not less than the maximum response value, then flushing

14   the data buffer to the stable storage device; and

15   if the expected delay is less than the maximum response value, then proceeding

16   to the step of determining the rate of incoming data.

1   3. The method of claim 1 wherein the step of determining whether a rate of incoming

2   data associated with the received transaction is higher than the current rate of flushing the

3   data buffer to the stable storage device comprises the steps of:

4   determining a recent rate of incoming transactions;

5   determining an amount of data submitted by an average transaction to the data

6   buffer;

7   computing a rate per second of an amount of data submitted to the data buffer by

8   the received transaction;

9   determining a target amount of data in the data buffer that is needed to achieve

10   sufficient batching in a flush operation based on the rate of incoming data and the flush

11   performance history;

12   determining an amount of data in the data buffer;

13   comparing the amount of data in the data buffer to the target amount of data; and

14   if the amount of data in the data buffer is not greater than the target amount of

15   data, then waiting for a next received transaction; and

16   if the amount of data in the data buffer is greater than the target amount of data,

17   then flushing the data buffer to the stable storage device.

1   4. The method of claim 1 wherein the step of flushing the data buffer to the stable

2   storage device comprises the steps of:

3   determining how many transactions are waiting for a flush of the data buffer;

4   determining an amount of data in the data buffer;

5   persisting the data in the data buffer to the stable storage device;

6   marking previously used memory by the data buffer as available for re-use;

7   resetting a data buffer size parameter to zero;

8   determining a duration for the writing, marking and resetting steps;

9   updating the data buffer flush performance history with the determined duration

10   for writing, marking and resetting;

11   determining an amount of data submitted by an average transaction to the data

12   buffer;

13   updating an average transaction data buffer size parameters; and

14   notifying waiting transactions that the flush has completed.

1   5. The method of claim 1 wherein the flush performance history is a table having a

2   plurality of buckets, each bucket having a minimum and a maximum amount of data

3   buffer data for which that bucket holds statistics, each bucket capable of storing a

4   plurality of entries, wherein each entry is an observed time to flush the data in the data

5   buffer to the stable storage device, wherein the step of updating the table of rates

6   comprises the steps of:

7   selecting a bucket in the plurality in which to record a data update based on the

8   amount of data in the data buffer; and

9   entering the data update into the selected bucket wherein the data update includes

10   a time duration for flushing the data in the data buffer to the stable storage device.

2   6. The method of claim 5 wherein the step of updating the table of rates further

3   comprises the steps of:

4   comparing the number of entries in the selected bucket to a bucket history length

5   constant; and

6   if the number of entries in the selected bucket is more than the bucket history

7   length constant, then removing the oldest entry.

1   7. The method of claim 3 wherein the flush performance history is a table having a

2   plurality of buckets, each bucket having a minimum and a maximum amount of data

3   buffer data for which that bucket holds statistics, each bucket capable of storing a

4   plurality of entries, wherein each entry is an observed time to flush the data buffer to the

5   stable storage device and wherein the step of determining a target amount of data

6   comprises the steps of:

7   a) selecting a first bucket in the flush performance history;

8   b) determining a minimum and maximum amount of data buffer data for which

9   the selected bucket holds flush statistics ;

10   c) based on the minimum and maximum amount of data, determining an average

11   amount of data in the data buffer for which the selected bucket holds flush performance

12   data;

13   d) calculating an average duration of a flush of the data buffer by averaging

14   values in the plurality of entries in the selected bucket;

15   e) calculating a number of times per second that a flush can be called based on the

16   average duration of a flush for the selected bucket;

17   f) calculating a rate of flushing for the selected bucket;

18   g) comparing the rate of flushing for the selected bucket to the rate of incoming

19   data;

20   h) if the rate of flushing for the selected bucket is not less than the rate of

21   incoming data, then taking the average data amount of data in the data buffer for the

22   selected bucket as the target amount of data;

23   i) if the rate of flushing for the selected bucket is less than the rate of incoming

24   data, determining whether there is at least one other bucket in the flush performance

25   history;

26   j) if there is at least one other bucket in the flush performance history, selecting a

27   second bucket in the flush performance history and repeating steps b) - j); and

28   k) if there is not at least one other bucket in the flush performance history, then

29   taking the average data amount of data in the data buffer for the selected bucket as the

30   target amount of data.

1   8. A method for efficiently batching transactions in a database system,

2   comprising the steps of:

3   receiving a transaction at the database system;

4   writing the received transaction data to a data buffer;

5   if the received transaction is an only transaction in the computerized

6   system, then flushing the data buffer to a database;

7   if the received transaction is not the only transaction in the computerized

8   system, then determining whether a maximum response time will be met for an oldest

9   transaction waiting for flush in the computerized system if the data buffer were flushed to

10   the database immediately;

11   if the maximum response time will not be met, then flushing the data buffer to the

12   database;

13   if the maximum response time will be met, then determining whether a rate of

14   data of incoming transactions is higher than a rate of flushing data from the data buffer to

15   the database if flushed immediately;

16   if the rate of data of incoming transactions is lower than the rate of flushing data

17   from the data buffer to the database if flushed immediately, then flushing the data buffer

18   to the database;

19   if the rate of incoming data is higher than the rate of flushing data from the data

20   buffer to the database if flushed immediately, then waiting to receive another transaction;

21   and

22   updating a flush performance history after each flush of the data buffer to the

23   database.

1   9. A transaction system that efficiently batches transactions, comprising:

2   a data interface to receive transactions;

3   a stable storage device to store processed transactions;

4   a memory including a data buffer and a system performance history, the system

5   performance history including system performance data about rates of flushing the data

6   buffer to the stable storage device and batch sizes of flushing the data buffer to the stable

7   storage device; and

8   a controller coupled to the interface, the stable storage device and the memory,

9   the controller configured to process the transactions, to write the processed transactions

10   to the data buffer, to flush the data buffer to the stable storage device according to the

11   system performance data from the system performance history, and to update the system

12   performance history with data from the flush of the data buffer to the stable storage

13   device.

1   10. The system of claim 9 wherein the controller further comprises:

2   a transaction processor to process transaction and to write the processed

3   transactions; and

4   a flush processor to flush the data buffer to the stable storage device and to update

5   the system performance history.

1   11. The system of claim 9 wherein the controller further comprises a maximum response

2   time value, the controller being further configured to determine whether a maximum

3   response time will be met for an oldest transaction in the data buffer and if the maximum

4   response time will not be met, to flush the data buffer to the stable storage device.

1   12. The system of claim 11 wherein the response time is determined based on

2   information from the system performance history

1   13. The system of claim 9 wherein the controller is configured to determine a rate of

2   incoming data and to determine a rate of flushing data, and if the rate of incoming data is

3   higher than the rate of flushing data, then to wait for further transactions in order to

4   develop a larger batch of processed transactions.

1   14. The system of claim 9 wherein the controller is further configured to adjust

2   frequency of flushing and amount of data per flush to the stable storage device based on

3   information from the system performance history.

1   15. The system of claim 9 wherein the system performance history is a table having a

2   plurality of buckets, each bucket having a minimum and a maximum amount of data

3   buffer data for which that bucket holds statistics, each bucket capable of storing a

4   plurality of entries, and wherein each entry is an observed time to flush the data in the

5   data buffer to the stable storage device.

1   16. The system of claim 15 wherein the controller is further configured to select a bucket

2   in the plurality of buckets in the system performance history into which to record a data

3   update, the controller to select based on an amount of data in the data buffer, the

4   controller is further configured to enter the data update into the selected bucket where the

5   data update includes a time duration for flushing data in the data buffer to the stable

6   storage device.

1   17. The system of claim 15 wherein the system performance history includes a bucket

2   history length constant, and wherein the controller is configured to compare the number

3   of entries in the selected bucket to the bucket history length constant and if the number of

4   entries exceeds the bucket history length constant, top remove old entries in excess of the

5   bucket history length constant.

1   18. A database system that efficiently batches transactions, comprising:

2   a data interface to receive database transactions;

3   a stable storage device having a database to store processed database transactions;

4   a memory including a data buffer and a system performance history, the system

5   performance history including system performance data about rates of flushing the data

6   buffer to the database and batch sizes of flushing the data buffer to the database; and

7   a controller coupled to the interface, the stable storage device and the memory,

8   the controller configured to process the transactions, to write the processed transactions

9   to the data buffer, to flush the data buffer to the database according to the system

10   performance data from the system performance history, and to update the system

11   performance history with data from the flush of the data buffer to the database.

1   19. A computer program product having a computer-readable medium including

2   computer program logic encoded thereon that, when performed on a computer system

3   directs the computer system to perform the method of:

4   receiving a transaction at the computerized system;

5   writing the received transaction data to a data buffer;

6   if the received transaction is an only transaction in the computerized

7   system, then flushing the data buffer to a stable storage device;

8   if the received transaction is not the only transaction in the computerized

9   system, then determining whether a maximum response time will be met for an oldest

10   transaction waiting for flush in the computerized system if the data buffer were flushed to

11   the stable storage device immediately;

12   if the maximum response time will not be met, then flushing the data buffer to the

13   stable storage device;

14   if the maximum response time will be met, then determining whether a rate of

15   data of incoming transactions is higher than a rate of flushing data from the data buffer to

16   the stable storage device if flushed immediately;

17   if the rate of data of incoming transactions is lower than the rate of flushing data

18   from the data buffer to the stable storage device if flushed immediately, then flushing the

19   data buffer to the stable storage device;

20   if the rate of incoming data is higher than the rate of flushing data from the data

21   buffer to the stable storage device if flushed immediately, then waiting to receive another

22   transaction; and

23   updating a flush performance history after each flush of the data buffer to the

24   stable storage device.

1   20. A system for efficiently batching transactions in a computerized system, comprising:

2   means for receiving a transaction at the computerized system;

3   means for writing the received transaction data to a data buffer;

4   means for flushing the data buffer to a stable storage device if the received

5   transaction is an only transaction in the computerized system;

6   means for determining whether a maximum response time will be met for an

7   oldest transaction waiting for flush in the computerized system if the data buffer were

8   flushed to the stable storage device immediately if the received transaction is not the only

9   transaction in the computerized system;

10   means for flushing the data buffer to the stable storage device if the maximum

11   response time will not be met;

12   means for determining whether a rate of data of incoming transactions is higher

13   than a rate of flushing data from the data buffer to the stable storage device if flushed

14   immediately if the maximum response time will be met;

15   means for flushing the data buffer to the stable storage device if the rate of data of

16   incoming transactions is lower than the rate of flushing data from the data buffer to the

17   stable storage device if flushed immediately;

18   means for waiting to receive another transaction if the rate of incoming data is

19   higher than the rate of flushing data from the data buffer to the stable storage device if

20   flushed immediately; and

21   means for updating a flush performance history after each flush of the data buffer

22   to the stable storage device.